

WHY COUNTING IS POLITICAL

1. Counting requires decisions about categorizing, about what (or whom) to include and exclude.
2. Measuring any phenomenon implicitly creates norms about how much is too little, too much, or just right.
3. Numbers can be ambiguous, and so leave room for political struggles to control their interpretation.
4. Numbers are used to tell stories, such as stories of decline ("we are approaching a crisis").
5. Numbers can create the illusion that a very complex and ambiguous phenomenon is simple, countable, and precisely defined.
6. Numbers can create political communities out of people who share some trait that has been counted.
7. Counting can aid negotiation and compromise, by making intangible qualities seem divisible.
8. Numbers, by seeming to be so precise, help bolster the authority of those who count.

ably reached the point of viability outside the womb, and so the state's interest had now gotten heavier and tipped the scale, or reached the "compelling point" in Blackmun's language. At this point, a state could prohibit abortion if it wished.

Like pregnancy, once a phenomenon has been converted into quantifiable units, it can be added, multiplied, divided, or subtracted, even though these operations have little meaning in reality. Numbers provide the comforting illusion that incommensurables can be weighed against each other, because arithmetic always "works": arithmetic yields answers. Numbers force a common denominator where there is none. They make it possible to reduce conflicts to a single dimension of size—big versus little, more versus less.

Finally, in our profoundly numerical contemporary culture, numbers are symbols of precision, accuracy, and objectivity. They suggest mechanical selection, dictated by the nature of the objects, even though

all counting involves judgment and discretion. By the time we are adults, the categorization part of counting is so much second nature that we tend to forget we do it. One has only to watch a child learning to count to see how self-conscious, learned, and precarious are the categorizations adults take for granted. Numerals hide all the difficult choices that go into a count. And certain kinds of numbers—big ones, ones with decimal points, ones that are not multiples of 10—not only conceal the underlying choices but seemingly advertise the prowess of the measurer, as if to say he or she could discriminate down to the gnat's knees. To offer one of these numbers is by itself a gesture of authority.

MAKING NUMBERS IN THE POLIS

Counting of any sort is a complex mental process, but measurement in the polis is a complex social process as well, nothing like counting up piles of peas and beans. Numbers in politics are measures of human activities, made by human beings, and intended to influence human behavior. They are subject to conscious and unconscious manipulation by the people being measured, the people making the measurements, and the people who interpret and use measures made by others. George Washington knew he had this kind of trouble on his hands in 1790, when the numbers from the first census began rolling in:

Returns of the Census have already been made from several of the States and a tolerably just estimate has been formed now in others, by which it appears that we shall hardly reach four millions; but one thing is certain: our real numbers will exceed, greatly, the official returns of them; because the religious scruples of some would not allow them to give in their lists; the fears of others that it was intended as the foundation of a tax induced them to conceal or diminished theirs; and thro' the indolence of the people, and the negligence of many of the Officers, numbers are omitted.¹¹

Measuring social phenomena differs from measuring physical objects because people, unlike rocks, respond to being measured. Try measuring people's heights and watch how they stretch their spines upward. Measurement is something like a mirror. When people know they are being measured, they try to imagine the results beforehand. They do a kind of stocktaking, whether of themselves or their program,

¹¹From J. C. Fitzpatrick, ed., *The Writings of George Washington* (Washington, D.C.: Government Printing Office, 1939) 31:329, cited in E. J. Kahn, Jr., *The American People* (Baltimore: Penguin Books, 1975), p. 45.

organization, or firm. Because measures always carry implicit norms, people wonder whether "their" measurements will meet the norms.

Measurement, like a mirror, triggers the natural desire to look good. People want to conform to their own ideals as well as to general social values. Measurement provokes people to "play the role" and to present themselves as they want to be seen. Such changes of behavior in response to being observed and measured are often called *reactive effects*, and reactivity is as pervasive among organizations as it is among individuals. Reactivity derives from two of the central features of the polis: influence and the interpretive nature of information. Since information is an instrument for shaping alliances and influencing others' behavior, it is no wonder that people try to control the impressions of themselves and their activities created by measures. And since people's choices are partly determined by other people's expectations and choices, it is no surprise that people try to give impressions of themselves as doing what they think other people want.

When something is measured, people tend to notice it more. Counting and measuring require looking and thinking, so the very process focuses people's attention on the things being measured. When the Head Start program was given a mandate to enroll handicapped children and required to submit enrollment figures, local directors started perceiving children as handicapped. One director noted, "We didn't know we had so many handicapped children until we started counting."¹² When the census added a category of ethnicity for Cajuns, the number of self-identified Cajuns shot up dramatically; as one commentator notes, "merely placing such an option on the census invites people to consider choosing it."¹³

The establishment of reporting and record-keeping mechanisms always brings out cases, as if the records exerted some kind of magnetic force on the things they record. Sometimes this happens because formal count normalizes a problem thought to be rare, and so encourages legitimizing something people were previously afraid or ashamed to discuss. This phenomenon is thought to be a major reason why reports of rape have escalated. Moreover, by creating groups, counting enables the victim of a stigmatized condition to come forward as a group member rather than as an individual. Record keeping also provides a check-

nel for reporting. Once an agency publicizes that it is keeping a count, people turn to that agency to report instances.

Record keeping is especially apt to stimulate reporting in the early phases of a policy issue. When a problem first gathers public attention, reformers begin to collect statistics and compare the present with the past. Every reform group needs to show that the problem it fights is not only big but somehow new and not something that has been tolerated for centuries. Reformers actively seek out cases, and because there is usually a relative lack of statistics for the previous period, the proportions of the problem loom large.

Many public problems are in fact things that have been tolerated for decades, if not centuries—alcoholism, child abuse, wife beating, elder abuse, environmental degradation, unequal access to health care, poor public schools, and congested cities, to name a few. Apparently dramatic growth rates in these problems reflect a decline in social tolerance of the phenomenon more than an increase in the phenomenon itself. As people's tolerance for any annoyance diminishes, their perception of the problem grows, just as annoying noises seem to get louder once we feel annoyed. Child abuse, some have argued, shows such a rapid rate of increase because the historically high level of violence toward children is no longer acceptable. People define child abuse more broadly, see it where they used to see parental prerogative and discipline, and report it because they can no longer tolerate it.¹⁴

If dissatisfaction induces counting, counting can also induce dissatisfaction and the desire for change. Stepping on the scale leads people to diets; surveying the wardrobe leads them to department stores; and balancing the checkbook leads them to the boss's office. When communities or organizations do a survey of their resources, the survey increases general awareness of problems and feeds demands for change. Conventional wisdom has it that the British National Health Insurance program got its critical stimulus from the army's induction exams for the Boer War at the turn of the century. The army found a high proportion of men unfit to serve, and the revelation prompted political leaders to invest in health care as a way of strengthening national manpower. Or, to take an American example, surveys of illiteracy and tests of students' knowledge almost always lead to curriculum reforms.

In public policy, measures are explicitly evaluative and are used to determine how people and organizations will be treated. Therefore, when people are measured in the polis, something more than their

¹² Robert Bogdan and Margaret Ksander, "Policy Data as a Social Process: A Qualitative Approach to Quantitative Data," *Human Organization* 39, no. 4 (1980): 304.

¹³ Lawrence Wright, "One Drop of Blood," *New Yorker*, July 25, 1994, pp. 46–55; quoted on p. 47.

Bogdan and Ksander, op. cit. (note 12), p. 304.

self-image and their desire to please is on the line: their fate is at stake. Some measures evaluate performance—of organizations, officials, employees, firms, or government agencies. They are used to determine rewards and punishments. Other measures evaluate needs—of individuals for social welfare benefits, of firms for small business loans or special tax treatment, of agencies for budget resources, for example. These measures are used to legitimize or dispute claims for resources and privileges.

People thus have a strong incentive to manipulate measures in the polis. When teachers' salaries are determined by how well their students perform on standardized tests, teachers start "teaching to the tests," focusing their entire efforts on boosting students' scores. When hospitals are judged by the cost per patient they incur, hospital administrators dump the difficult-to-treat and uninsured patients onto other hospitals. When job training programs are evaluated by the proportion of trainees they place in jobs, they tend to select the most employable people as trainees rather than the people who most need training, and they eagerly place trainees in easy-to-find, dead-end jobs. In all these examples, no one actually falsifies the numbers. Rather, people change their behavior in response to being measured, and the changed behavior leads to results that are actually different from what they would have been without the measure.

Understanding reactivity is so important in policy analysis because unlike deliberate falsification of numbers, it is an inextricable feature of social measurement.¹⁵ Moreover, reactivity violates the canons of good scientific practice, on which all statistical reasoning is based. Scientific method assumes a strict separation between the observer and the observed. The subject of measurement should never measure himself or herself. To do so would be the essence of subjectivity, and scientific measurement pursues the ideal of objectivity, where neither the subject nor the observer has incentive or opportunity to manipulate the way a measure appears. In experimental research involving human subjects, scientists use the term "blindness" to mean lack of knowledge about whether the subject is in the control group or the experimental

group. In a single-blind experiment, the subject does not know which group he or she is in, but the observer does. In a double-blind experiment, a third party codes the subjects so that neither the subject nor the observer (measurer) knows. Double-blindness is thought to make for the most accurate, unbiased, nonreactive measurement. Measurement in the polis is more appropriately characterized by heightened vision on the part of subjects and measurers rather than by any kind of blindness.

Separation between the people being measured and the people doing the measuring is rarely, if ever, possible in politics. The vast majority of measures in policy debates are official statistics, gathered, analyzed, and reported by the agencies whose performance and needs they assess. Crime statistics are produced by crime-fighting agencies and housing statistics by housing development agencies. Thus, the measures often reflect as much on the behavior of the measurers as the measured, and measurers, too, have an incentive to manipulate.

The clearance rate, for example, is supposed to measure police effectiveness in solving crime. It is loosely defined as the percentage of crimes known to the police that police believe they have solved. The clearance rate can be raised either by solving more crimes (increasing the numerator) or reducing the total number of crimes known to the police (decreasing the denominator). Police, it turns out, have a fair amount of discretion in classifying complaints as "unfounded" or "suspicious," and these complaints are not entered as "crimes known to police." A man reports a burglary, for example, but police can't find physical evidence of a break-in and suspect the man really wants to claim a loss to his insurance company. Since police performance looks better the lower the total crimes and the higher the clearance rate, there are subtle pressures for them to regard reports as unfounded or suspicious. Thus, because the clearance rate is used as an evaluative measure, it influences the way police count.¹⁶

Even when the measurer is organizationally separate from the measured, such as when an accounting firm audits a business firm or one government agency monitors another, the separation is only nominal. The very fact that one organization is measuring another links them in a relationship of cooperation and influence. The accounting firm needs the business of the firms it audits, and so must be responsive to its client firms' desires; and it is still dependent on the data provided

¹⁵One very influential, if not predominant, school of thought in social science aspires to eliminate reactivity in social measurement and thus to make the social sciences meet the presumed standards of the physical sciences. (In fact, measurement in the physical sciences is not devoid of some of these problems.) The classic work in this tradition is Eugene J. Webb et al., *Unobtrusive Measures: Nonreactive Research in the Social Sciences* (Chicago: Rand McNally, 1966). I argue the folly of such a dream in this chapter.

¹⁶Jerome Skolnick, *Justice Without Trial: Law Enforcement in Democratic Society* (New York: Wiley, 1975), chap. 8.

by client firms. The government monitoring agency needs the cooperation of the agencies it monitors, both because it needs data and because the people in the monitoring agency would rather have pleasant working relationships than hostile ones. Measurers need the people they measure.

Measurers also have power over the fate of the measured, since measuring is done to help decide on policy actions. The census taker not only counts heads, but also determines the apportionment of seats in the House of Representatives and the allocation of federal revenue-sharing funds. The health inspector makes bacteria counts, but with them closes down restaurants and destroys reputations. The Antitrust Division of the Justice Department measures market concentration to decide whether to allow mergers and acquisitions. Because policy measurement is always linked to benefits and penalties, the measured try to influence the measurers, occasionally with outright bribes, but more often with pleading, cajoling, and selective disclosure.

Evaluative measurement sometimes makes strange bedfellows. A suspect who confesses to many crimes is more valuable to police than a suspect who confesses to only one, because the multiple offender helps "clear" several crimes at once. Police are more willing to bargain with a suspect who can potentially raise the clearance rate, by offering reduced charges and counts, arranging for only one of the offenses to go on the record, and promising immunity from prosecution for the other crimes. The use of the clearance rate as an evaluative measure thus actually reverses the hierarchy of punishments in the criminal justice system in that the multiple offender is treated with relatively more leniency than the single offender.¹⁷

A similar logic seems to have supported the Justice Department's decision not to prosecute E. F. Hutton executives when the brokerage firm was caught in an elaborate check overdraft scheme. According to the *New York Times*, the Justice Department "contended that Hutton's guilty pleas to 2,000 counts of mail and wire fraud were far more valuable to the interests of justice, including deterring any such practices by other companies, than would have [been] the prosecution of individuals in a trial that might have gone on for many months."¹⁸

The check-writing scam illustrates how, in a measuring relationship, the thing being counted can become a commodity that is strategically traded between the measurer and the measured. It is bad to

commit a crime, and committing a large number is worse than committing a small number; but since guilty pleas are counted in the justice system as a measure of crime-fighting success, crimes become bargaining chips, and it is better for the criminal to have a lot of crimes to confess than a few. At different points in the investigation of E. F. Hutton's behavior, its executives had strong motivations first to conceal the number of check overdrafts and later to reveal as many as possible. On the other side, the incentives for the Justice Department were to collect a lot of guilty pleas for trivial crimes rather than to prosecute and seek to change a large, systematic pattern of crime.

This dynamic in which the thing measured becomes a strategically traded commodity occurs whenever the measurer is rewarded or punished for finding instances of the thing measured. All regulatory agencies are in this position. They monitor the behavior of regulated firms by counting violations. They may be encouraged and rewarded by Congress (with higher budget appropriations) for finding violations, or punished by Congress in a period of antiregulatory sentiment. Regulators may be rewarded by regulated firms with lucrative jobs for not finding too many violations, or for devising sophisticated ways of "counting as" that minimize violations.

A more conspicuous aspect of number manipulation in the polis occurs in the selection of measures from the vast range of possibilities. (Remember my elephant). If, as the saying goes, where one stands depends on where one sits, so does how one counts. In election campaigns, incumbents and challengers choose different measures of inflation or poverty or unemployment. One side wants measures to make these perennial problems look low and declining, while the other wants to make them look high and growing.

Congress and the defense agencies play cat-and-mouse games with measures of cost. The cost of the F-14 fighter plane (the Tomcat) has been figured at anywhere from \$26 million to \$35 million per plane. One measure is program cost, or the total cost for developing and building the planes divided by the number of planes actually built. That gives a very high cost per plane. Another measure is marginal cost, or the cost of building the next plane, assuming production is already running. That gives a lower cost per plane. Yet another measure is the hypothetical cost for a large package deal of, say, 1,000 or 1,500 planes, which turns out to give the lowest cost of the three measures.¹⁹ I leave it to readers to guess which measure might have been

¹⁷Ibid.

¹⁸Nathaniel C. Nash, "Rep. Hughes Criticizes Hutton," *New York Times*, July 12, 1985, p. D5.

¹⁹The example comes from James Fallows, *National Defense* (New York: Vintage Books, 1981), p. 37.

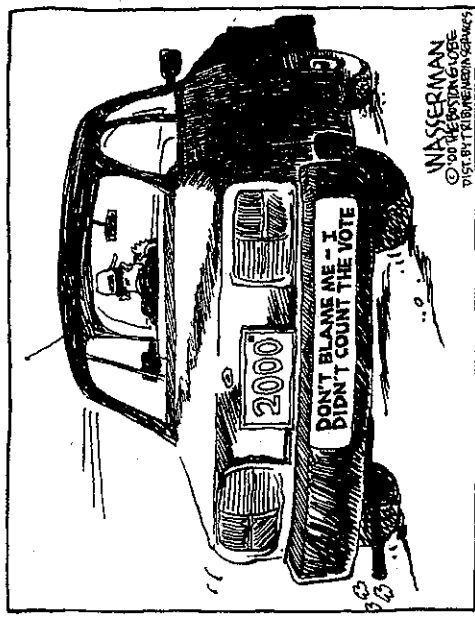
touted by the Air Force as it sought original funding from Congress, which one was bandied about as Congress reconsidered the budget midway through the building program, and which one was thrown back at the Air Force by an angry Congress years later.

People in their role as fund-raisers, either as philanthropic leaders or as heads of agencies, always use measures that make "their" problems seem as big as possible. But once an agency has been in operation, allegedly working to solve the problem, its leaders also need to show that the problem has diminished, so that the agency will look effective. Thus, agencies might choose absolute number of cases as a measure of a problem early in the life of a policy issue, but proportion of known cases treated in later stages of an issue. And, agencies are very likely to present different measures of the problem to different audiences. For example, an environmental regulatory agency might count total sources of pollution to justify its budget, but count number of sources closed or number of disciplinary actions to show its own effectiveness.

Business firms, too, select measures to show their performance in a certain light. They want to look highly profitable to investors, but not so profitable to the Internal Revenue Service. Firms can manipulate the value of their assets by allocating depreciation. They can change their cost and productivity figures by how they allocate overhead and other fixed costs to individual products. If they have lost big environmental or consumer product liability suits, they can tally up estimated future payments, guessing how many suits they might lose, in order to wipe out their assets altogether. With the help of its accountants, Johns-Manville, a large asbestos manufacturer, used this trick to show bankruptcy, despite its ranking as number 181 in the Fortune 500 and its assets of over \$2 billion.²⁰

Measures tend to imply certain solutions to a problem, so people who have particular solutions to peddle will promote the measures that point to their solutions. Social workers benefit from such definitions of poverty as the number of broken families (another metaphor) or the number of families with children at risk of abuse. The medical industry is wont to measure poverty as the number of people who cannot afford its services, the so-called medically indigent. Building contractors benefit more from a definition of poverty based on housing such as the number of families who live in substandard dwellings. Educators tout measures of illiteracy as indicators of poverty. Farmers with surplus agricultural products might like a definition based on malnutrition. Economists, to their credit, are among the few who mea-

²⁰ Paul Brodeur, "The Asbestos Industry on Trial," *New Yorker*, June 10, 1985, p. 49.

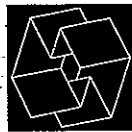


sure poverty as the number of people lacking in income. Measures, in short, obey the law of the hammer: Give a small boy a hammer and he'll find plenty of things that need pounding.²¹

Measures in the polis are not only strategically selected but strategically presented as well. Numbers never stand by themselves in policy debates; they are clothed in words and symbols and carried in narrative stories. Articles and speeches rarely say simply that a rate changed from this to that number. They give stage directions, telling the audience that rates "rose rapidly," "dropped substantially," "moved sluggishly," or otherwise behaved dramatically. Without these cues, we would not know how to react to the naked numbers.

News stories about employment data provide great examples of number costuming, because unemployment is a double-edged sword. (It's good for unemployment to be low so fewer people need assistance, but bad for unemployment to be too low—any lower than 5 percent economists deem inflationary.) Because unemployment figures had been hovering just above 5 percent for the first half of 1996, policymakers awaited the August reports with trepidation. Would the unemployment rate trigger the Federal Reserve to raise interest rates again? When the news came, unemployment had risen from 5.3 percent in June to 5.4 percent in July, only a tenth of a percent. The economy had added 193,000 new jobs in July, "much more tame than June's huge gain of 220,000," said the *Wall Street Journal* with an audible

²¹ This law probably has folk origins, but I learned it from Abraham Kaplan, *The Conduct of Inquiry* (New York: Harper & Row, 1968).



NUMERICAL STRATEGIES IN PROBLEM DEFINITION

1. People react to being counted or measured, and try to "look good" on the measure.
2. The process of counting something makes people notice it more, and record keeping stimulates reporting.
3. Counting can be used to stimulate public demands for change.
4. When measurement is explicitly used to evaluate performance, the people being evaluated try to manipulate their "scores."
5. The power to measure is the power to control. Measurers have a lot of discretion in their choice of what and how to measure.
6. Measuring creates alliances between the measurers and the measured.
7. Numbers don't speak for themselves, and people try to control how others will interpret numbers.

sigh of relief. (Neither June's "huge" gain nor the average monthly gain of 237,000 jobs in the first half year probably offered much solace to the 7 million people out of work.) The story's lead played the numbers as a melodrama: "Somewhat subdued job growth and declining hourly wages in July offered the clearest signs yet that the economy is throttling down from the second quarter's torrid pace." Torrid pace. This was the same week Congress passed welfare reform, and directly above the story about the "benign" new unemployment data was one about state governors "in a cold sweat" over "the unenviable task of finding full-time jobs for two million of the least skilled, least employable Americans."²²

²² "Economy Shows Signs of Slower Pace; Benign Employment Data Lift Markets," and "Welfare Law's Work Rules Worry States," both in *Wall Street Journal*, August 5, 1996, p. A2.

Numbers in policy debates cannot be understood without probing how they are produced by people: What makes people decide to count something and then find instances of it? How are the measurers and the measured linked together? What incentives do people have to make the numbers appear high or low? And what opportunities do they have to behave strategically? People change the activities that are being measured. They try to influence the measurers. The exercise of counting makes them notice things more. Measurers change the way they count because their measures affect how they, not only the measured, are treated. The things being counted become bargaining chips in a strategic relationship between the measurers and the measured, so that at different points in the relationship, there are very different pressures to reveal or conceal. The choice of measures is part of strategic problem definition, and the results of measures take on their political character only with the costume of interpretive language.

If numbers are thus artifacts of political life, and if they are themselves metaphors, symbols, and stories, are they "real" in any sense? Numbers are always descriptions of the world, and as descriptions, they are no more real than the visions of poems or paintings. Their vision of experience may correspond more or less with popular visions, just as realist, impressionist, and abstract expressionist paintings correspond more or less with common visions. Numbers are real as artifacts, just as poems and paintings are artifacts that people collect, recite, display, and respond to. But the dominance of numbers as a mode of describing society in public policy discussions is only a recent, and perhaps temporary, phenomenon in cultural history—not the result of some underlying reality of numbers.